

RESEARCH LETTER

Cross-correlating soil aggregate stability methods to facilitate universal interpretation

Deborah Aller  | Joseph P. Amsili  | Harold M. van Es 

Section of Soil and Crop Sciences, School of Integrative Plant Science, Cornell University, Ithaca, New York, USA

Correspondence

Deborah Aller, Section of Soil and Crop Sciences, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA.

Email: da352@cornell.edu

Assigned to Associate Editor Veronica Acosta-Martinez.

Funding information

New York State Environmental Protection Fund; Foundation for Food and Agricultural Research, Grant/Award Number: 523926; General Mills; The Samuel Roberts Noble Foundation

Abstract

Aggregate stability is a critical physical indicator of soil health. However, multiple methods are used for measuring aggregate stability, making it difficult to compare results and limiting universal interpretations in soil health assessment frameworks like Soil Health Assessment Protocol and Evaluation. We cross-correlated three common water-stable aggregate methods (WSA_{CASH} , WSA_{ARS} , and WSA_{SLAKES}) using a dataset of nearly 1400 samples and developed pedotransfer functions using random forest models to evaluate method performance. We found that the WSA_{ARS} and WSA_{CASH} methods can be reasonably cross correlated through pedotransfer functions because they use similar processes for estimating aggregate strength. Conversely, the WSA_{ARS} and WSA_{SLAKES} methods are not transferable. We suggest that the WSA_{ARS} aggregate stability method is the most established and best reference method for use in soil health analysis frameworks. Interpretation consistency will lead to more robust comparisons of aggregate stability as a key physical soil health indicator.

1 | INTRODUCTION

Healthy soils are vital for sustainable agricultural production and environmental protection. Good soil structure, which depends on the presence of stable aggregates, is crucial for supporting the numerous processes and functions provided by a healthy soil (Amézqueta, 1999). Soil aggregation leads to improved porosity, infiltration, root proliferation, and water availability for plants, and helps to resist soil erosion, runoff, and compaction (Kemper & Rosenau, 1986). Stable aggregates also aid in physical protection of soil organic matter and creation of habitat for soil organisms (Six & Paustian, 2014).

Abbreviations: CASH, Comprehensive Assessment of Soil Health; MAP, mean annual precipitation; MAT, mean annual temperature; SHAPE, Soil Health Assessment Protocol and Evaluation; SMAF, Soil Management Assessment Framework; WSA, water-stable aggregates.

Aggregate stability generally refers to the ability of soil aggregates to withstand wetting and the impact of rain drops (Amézqueta, 1999; Moebius et al., 2007; Topp et al., 1997). Decades ago, it was proposed as a physical indicator of soil quality and remains as the most common laboratory measured physical soil health indicator (Arshad & Coen, 1992). Today, current US soil health assessment frameworks, namely, the Soil Management Assessment Framework (SMAF; Andrews et al., 2004), Comprehensive Assessment of Soil Health (CASH; Moebius-Clune et al., 2016), and Soil Health Assessment Protocol and Evaluation (SHAPE; Nunes et al., 2021), as well as commercial laboratories, recognize aggregate stability as a key physical indicator of soil health. However, consistency in methods used when monitoring changes in soil health over time is important (Rieke et al., 2022). Thus, having multiple commonly used methods for

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Agricultural & Environmental Letters* published by Wiley Periodicals LLC on behalf of American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America.

measuring aggregate stability creates challenges when comparing results.

Currently, three protocols for measuring percentage water-stable aggregates (WSA) are commonly used: the Cornell sprinkle infiltrometer method of the CASH framework (WSA_{CASH} ; Moebius-Clune et al., 2016), the wet sieve procedure that was originally developed by the USDA-ARS (WSA_{ARS} ; Kemper & Rosenau, 1986; Yoder, 1936), and the SLAKES smart phone app (WSA_{SLAKES} ; Fajardo et al., 2016). All aggregate stability methods can capture changes in soil structure but are not equally powerful in detecting management effects (Rieke et al., 2022). WSA_{SLAKES} is a new methodology with thus far limited use but is otherwise advantageous in terms of cost, availability, and ease of use (Flynn et al., 2020). The WSA_{ARS} method has a long history of use in soil science research and traditionally been the most widely reported. WSA_{CASH} is generally more robust, less sensitive to changes and uncertainties, with detecting agronomic management effects than WSA_{ARS} , while WSA_{SLAKES} has shown, in some cases, results that were inconsistent with the other methods (Rieke et al., 2022). Although the WSA_{CASH} approach is used by some university soil health laboratories and has the most samples processed in recent years, the method may be less attractive for commercial expansion due to limited equipment availability.

The SHAPE interpretation tool expands the SMAF and CASH frameworks by using a US nationwide dataset of commonly measured soil health indicators to provide a quantitative and interpretive soil health score based on soil peer groups. Scoring curves are provided by Bayesian linear regression modeling based on soil suborder and texture class, while adjusting for mean annual temperature (MAT) and precipitation, which knowingly impact soil properties (Nunes et al., 2021, 2024). The current SHAPE dashboard (https://paparker.shinyapps.io/shape_app/) provides aggregate stability scores for two different percentage WSA methods: WSA_{CASH} and the method used by the Kellogg Soil Survey Laboratory (WSA_{KSSL} ; Kellogg Soil Survey Laboratory wet sieve procedure; Soil Survey Staff, 2022), a method that is similar to WSA_{ARS} . It is proposed that one is the reference method for soil health scoring functions so results are comparable. Other methods then need to be cross-correlated with the reference method. We think that the longer history of the WSA_{ARS} method makes it most suitable to serve as a reference standard in soil health analyses and other methods should be cross-correlated to it.

Our objectives were to (1) evaluate cross-correlations between three currently used aggregate stability methods and (2) develop pedotransfer functions using random forest (RF) modeling to predict WSA_{ARS} , the proposed reference method, from WSA_{CASH} , WSA_{SLAKES} , and other soil health indicator data for use in scoring frameworks such as SHAPE.

Core Ideas

- Different approaches for measuring aggregate stability prevent generalized result interpretation.
- The water-stable aggregate wet sieve procedure (WSA_{ARS}) is proposed as the reference method for interpretation.
- Other soil aggregate stability methods can be variably correlated with WSA_{ARS} .

2 | MATERIALS AND METHODS

The data used were from the North American Project to Evaluate Soil Health Measurements (NAPESHM) conducted by the Soil Health Institute, which sampled 2012 experimental units from 688 replicated treatments located at 124 long-term experimental agricultural research sites across the United States, Mexico, and Canada (Norris et al., 2020). For this analysis, data from Canada and Mexico were removed because the initial purpose of this project was to cross-correlate WSA methods for US soils. Data points that did not include all aggregate stability methods were also removed. The final dataset contained 1399 samples (0- to 15-cm depth) from 87 long-term experiments in the contiguous United States.

The WSA methods evaluated were WSA_{CASH} (Moebius-Clune et al., 2016), WSA_{ARS} (Kemper & Rosenau, 1986; Yoder, 1936), and WSA_{SLAKES} (Fajardo et al., 2016). A detailed description of each method is provided elsewhere (Rieke et al., 2022). In brief, WSA_{ARS} and WSA_{CASH} both evaluate aggregates on top of 0.25-mm sieves. WSA_{ARS} places aggregates 1–2 mm in diameter on a 0.25-mm sieve and applies external energy by standard oscillations while submerged in deionized water to induce slaking of unstable aggregates, whereas WSA_{CASH} spreads aggregates 0.25–2 mm in diameter on a 0.25-mm sieve and uses rainfall energy (2.5 J over 5 min) from a rainfall simulator. WSA_{SLAKES} does not apply external energy to disperse aggregates but instead uses an image recognition software to measure the disintegration (slaking) of three air-dried aggregates following a 10-min submersion in shallow water. The WSA_{SLAKES} data were normalized by multiplying by 100 to be on the same scale as the WSA_{ARS} and WSA_{CASH} data.

2.1 | Statistical analysis

All data analyses were performed in R Version 4.2.1 (R Core Team, 2022). Quadratic regression modeling among WSA methods was carried out, using the `lm` function, across coarse (sand, loamy sand, and sandy loam), medium (loam, silt loam,

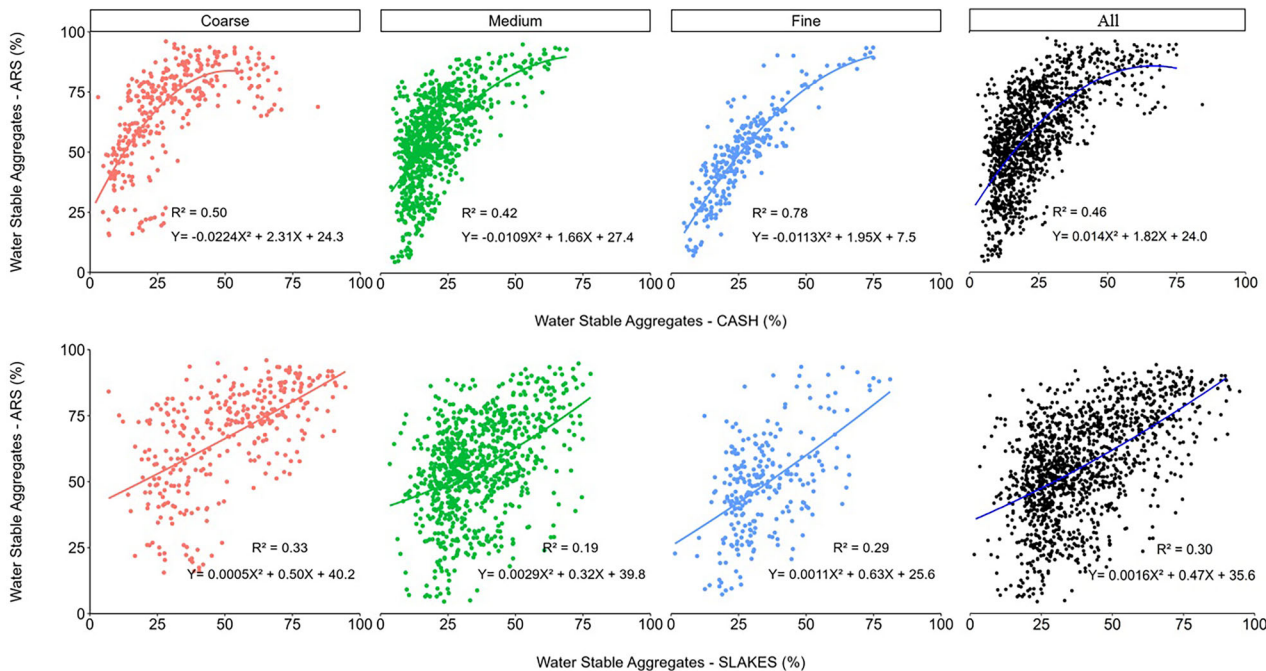


FIGURE 1 Quadratic regression models across coarse, medium, and fine soil texture groups, and all points for WSA_{CASH} versus WSA_{ARS} (top panel) and WSA_{SLAKES} versus WSA_{ARS} (bottom panel) aggregate stability methods. WSA, water-stable aggregates.

sandy clay loam, and silt), and fine (clay loam, silty clay loam, silty clay, sandy clay, and clay) soil texture groups. Three RF models were created to predict WSA_{ARS} using the randomForest package, which implements Breiman’s RF algorithm (Breiman, 2001; Liaw & Weiner, 2022). The out-of-bag (OOB) root mean squared error (RMSE) was used to evaluate and validate the RF model as it estimates the prediction error of the bagged model and requires no cross-validation. One RF model utilized only soil and climate predictor variables: sand, silt, clay, soil organic carbon, MAT, and mean annual precipitation (MAP). Two other RF models additionally included either WSA_{CASH} or WSA_{SLAKES}. Variable importance metrics (mean square error [MSE] increase and node purity increase) were plotted using the randomForestExplainer package (Palusynska et al., 2022). The RF models were compared to analogous multiple linear regression models (using the lm function) containing the three most important variables identified by RF models, which were soil organic carbon, MAP, and clay, and either no WSA, WSA_{CASH}, or WSA_{SLAKES}.

3 | RESULTS AND DISCUSSION

A comparison of aggregate stability across three soil texture groups revealed that the WSA_{CASH} method on average led to lower percentage stability values compared to the WSA_{ARS} method, that is, 37.5%, 34.9%, and 22.5% lower for coarse, medium, and fine texture groups, respectively, similar

to Van Eerd et al. (2018). The WSA_{SLAKES} method averaged intermediate to the two other methods. All displayed the highest average values for coarse-textured soils compared to medium- and fine-textured soils, but this was most dramatic for WSA_{ARS} and WSA_{SLAKES}. For WSA_{SLAKES}, this is likely attributed to sand particles rapidly falling out of suspension following aggregate submersion in water (Rieke et al., 2022).

Quadratic regression analysis showed that WSA_{ARS}-WSA_{CASH} were a better model fit ($R^2 = 0.46$) and had lower model uncertainty compared to WSA_{ARS}-WSA_{SLAKES} ($R^2 = 0.30$). For coarse- and medium-textured soils, there can be a wide range of values for WSA_{ARS} when WSA_{CASH} values are low, but the range of WSA_{ARS} values gets narrower as WSA_{CASH} gets higher (above ~30%), that is, WSA_{ARS} shows greater sensitivity with low-stability aggregates, while WSA_{CASH} is more discerning with high-stability aggregates. The relationship between WSA_{ARS}-WSA_{CASH} had the tightest fit for fine textured soils ($R^2 = 0.78$), followed by coarse soils ($R^2 = 0.50$), and then medium textured soils ($R^2 = 0.42$). We conclude that the WSA_{ARS} and WSA_{CASH} can reasonably be cross-correlated through texture group specific regression equations.

The overall poor fit between WSA_{ARS} and WSA_{SLAKES} was also observed when separating texture groups (Figure 1), and quadratic regression models did not improve the fit compared to linear regression. The weaker correlations between WSA_{ARS} and WSA_{SLAKES} compared to WSA_{ARS} and WSA_{CASH} may reflect the more substantial difference in the measurement process (external energy with WSA_{ARS}

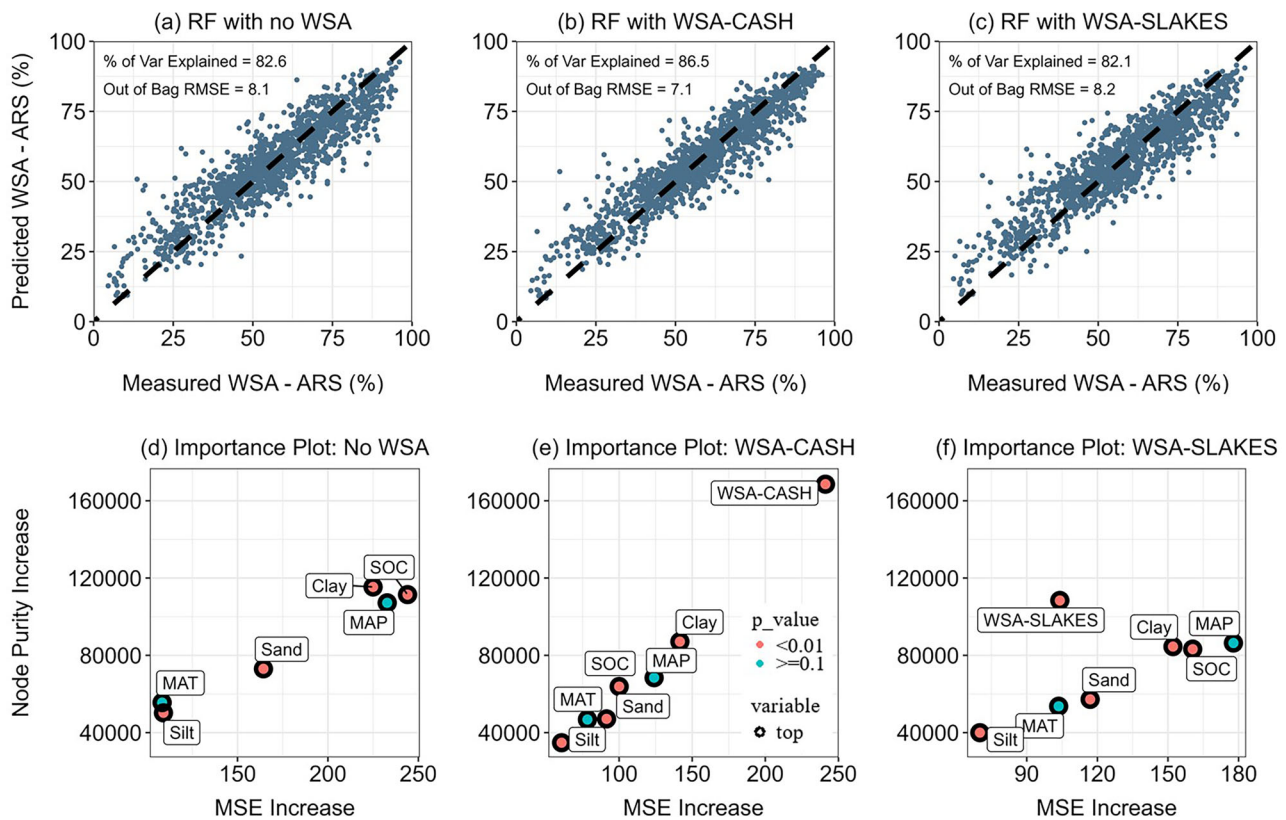


FIGURE 2 Measured versus predicted water-stable aggregates (WSA_{ARS}) for three random forest (RF) models (a–c) and corresponding multiway variable importance plots (d–f). All RF models utilized the predictor variables of sand, silt, clay, soil organic carbon, mean annual temperature (MAT) and mean annual precipitation (MAP), and either without WSA (a and d), with WSA_{CASH} (b and e) or with WSA_{SLAKES} (c and f). Validation metrics within plots (a–c) include percentage variance explained and out of bag root mean square error (RMSE). In variable importance plots (d–f), mean square error (MSE) increase ranks variables by how much they reduce the MSE, where higher MSE increase values imply that if the variable was removed from the model, it would lead to a larger increase in MSE. Node purity increase ranks variables by how well they decrease the residual sum of squares across all trees when they are used as a splitting variable.

[submerged shaking] and WSA_{CASH} [raindrops] vs. none with WSA_{SLAKES}). Rieke et al. (2022) also found higher variability with WSA_{SLAKES} , and agronomic management effects on WSA_{SLAKES} were at times inconsistent with the other methods. We conclude that the WSA_{ARS} and WSA_{SLAKES} cannot be readily cross correlated through texture group specific regression equations.

The RF model that is solely based on soil data with no WSA measurements explained 82.6% of the variance with an OOB-RMSE of 8.1 (Figure 2a). The model that also included the WSA_{CASH} (Figure 2b) improved predictability of WSA_{ARS} (OOB RMSE = 7.1; Figure 2b). The RF model that included WSA_{SLAKES} with the soil and climate data (Figure 2c) did not improve predictability and in fact slightly reduced it from the soil and climate data-only model. This indicates that the RF model estimation of the reference method WSA_{ARS} can be reasonably done with basic soil and climate data and additional WSA_{CASH} data improve precision. But additional WSA_{SLAKES} data showed no meaningful improvement over the RF model without WSA data. The RF models reduced

the RMSE by 62.4%–77.9% compared to analogous multiple linear regression models (models not shown), indicating that RF models greatly improved predictability compared to traditional regression techniques.

The RF variable importance plot for the WSA_{ARS} prediction model with WSA_{CASH} revealed that WSA_{CASH} was the most prominent predictor variable in the model for reducing the MSE (Figure 2e). However, WSA_{SLAKES} was tied with MAT as being the three least most important variables for reducing the MSE (Figure 2f). MAP, soil organic carbon, clay, and sand were all more important for prediction of WSA_{ARS} compared to WSA_{SLAKES} . This implies that WSA_{SLAKES} results do not add value to the estimation of WSA_{ARS} beyond the soil and climate variables that were included. Conversely, WSA_{CASH} measurements can improve WSA_{ARS} estimates because of the additional variables. The three RF models are available for download at the Harvard Dataverse (Amsili et al., 2024).

In all, we think that the WSA_{ARS} approach is the most established and widely researched method for soil aggregate

stability and therefore best serves as a reference method for use in interpretation frameworks like SHAPE (Nunes et al., 2024). The WSA_{CASH} rainfall simulation approach, while an equally or more robust method, may not be readily scaled up across laboratories due to limited availability of the apparatus. However, WSA_{CASH} results can be readily converted into WSA_{ARS} values by using RF models. Prediction of WSA_{ARS} from WSA_{SLAKES} is not recommended since WSA_{SLAKES} is a poorer predictor variable of WSA_{ARS} compared to WSA_{CASH} (Figures 1 and 2). The different assessment process with WSA_{SLAKES} makes it less comparable with more established methods, although its assessment approach may have value of its own.

4 | CONCLUSION

Soil health assessment frameworks are helping farmers move toward adopting more regenerative agriculture practices. SHAPE is a leading tool to provide regionally relevant soil health interpretations across the contiguous United States. However, multiple methods for measuring aggregate stability have created inconsistent results and prevented universal interpretations. We used a subset of the NAPESHM dataset to compare three widely used water-stable aggregate methods: WSA_{CASH} , WSA_{ARS} , and WSA_{SLAKES} . Primarily due to its historical use, we think that the WSA_{ARS} aggregate stability method is the most appropriate reference for use in the SHAPE scoring curves. WSA_{CASH} is a robust method and results can be cross-correlated and translated to WSA_{ARS} values. But WSA_{SLAKES} results are not as robust and should be interpreted using its own standards. Having consistency with methods will strengthen monitoring of soil health changes over time and interpretations for farmers and other agriculture stakeholders.

AUTHOR CONTRIBUTIONS

Deborah Aller: Writing—original draft; writing—review and editing. **Joseph P. Amsili:** Conceptualization; data curation; formal analysis; writing—review and editing. **Harold M. van Es:** Conceptualization; funding acquisition; writing—review and editing.

ACKNOWLEDGMENTS

This research was supported by a contract with the New York State Environmental Protection Fund, which was administered by the New York State Department of Agriculture and Markets. The authors thank the reviewers for their valuable feedback that improved the quality of this manuscript. We acknowledge Dr. Kristen Veum of USDA-ARS for collaborating on the SHAPE project. We also thank colleagues who have made their long-term research sites available for the North American Project to Evaluate Soil Health Mea-

surements (NAPESHM) project, as well as staff of the Soil Health Institute and laboratories that conducted the sampling and analyses. The NAPESHM project is part of a broader effort titled, “Assessing and Expanding Soil Health for Production, Economic, and Environmental Benefits.” The project was funded by the Foundation for Food and Agricultural Research (grant ID 523926), General Mills, and The Samuel Roberts Noble Foundation. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the Foundation for Food and Agriculture Research.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Deborah Aller  <https://orcid.org/0000-0002-2006-6985>

Joseph P. Amsili  <https://orcid.org/0000-0002-8293-5824>

Harold M. van Es  <https://orcid.org/0000-0001-9822-9476>

REFERENCES

- Amézketa, E. (1999). Soil aggregate stability: A review. *Journal of Sustainable Agriculture*, 14(2–3), 83–151. https://doi.org/10.1300/J064v14n02_08
- Amsili, J. P., Aller, D., & van Es, H. M. (2024). *Random forest models from “cross-correlating soil aggregate stability methods to facilitate universal interpretation.”* Harvard Dataverse. <https://doi.org/10.7910/DVN/LEBTIV>
- Andrews, S. S., Karlen, D. L., & Cambardella, C. A. (2004). The soil management assessment framework: A quantitative soil quality evaluation method. *Soil Science Society of America Journal*, 68(6), 1945–1962. <https://doi.org/10.2136/sssaj2004.1945>
- Arshad, M. A., & Coen, G. M. (1992). Characterization of soil quality: Physical and chemical criteria. *American Journal of Alternative Agriculture*, 7, 25–31. <http://www.jstor.org/stable/44479646>
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Fajardo, M., Mcbratney, A. B., Field, D. J., & Minasny, B. (2016). Soil slaking assessment using image recognition. *Soil and Tillage Research*, 163, 119–129. <https://doi.org/10.1016/j.still.2016.05.018>
- Flynn, K. D., Bagnall, D. K., & Morgan, C. L. S. (2020). Evaluation of SLAKES, a smartphone application for quantifying aggregate stability, in high-clay soils. *Soil Science Society of America Journal*, 84(2), 345–353. <https://doi.org/10.1002/saj2.20012>
- Kemper, W., & Rosenau, R. (1986). Aggregate stability and size distribution. In D. Sparks (Ed.), *Methods of soil analysis. Part 1—Physical and mineralogical methods* (pp. 425–442). Soil Science Society of America.
- Liaw, A., & Weiner, M. (2022). Randomforest: Breiman and Cutler’s Random Forests for Classification and Regression (R package version 4.7-1.1) [Computer software]. CRAN. <https://cran.r-project.org/web/packages/randomForest/index.html>
- Moebius, B. N., Van Es, H. M., Schindelbeck, R. R., Idowu, O. J., Clune, D. J., & Thies, J. E. (2007). Evaluation of laboratory-measured soil

- properties as indicators of soil physical quality. *Soil Science*, 172, 895–912. <https://doi.org/10.1097/SS.0B013E318154B520>
- Moebius-Clune, B. N., Moebius-Clune, D. J., Gugino, B. K., Idowu, O. J., Schindelbeck, R. R., Ristow, A. J., van Es, H. M., Thies, J. E., Shayler, H. A., McBride, M. B., Kurtz, K. S. M., Wolfe, D. W., & Abawi, G. S. (2016). *Comprehensive assessment of soil health—The Cornell framework manual* (3rd ed.). Cornell University. <https://soilhealth.cals.cornell.edu/manual/>
- Norris, C. E., Bean, G. M., Cappellazzi, S. B., Cope, M., Greub, K. L. H., Liptzin, D., Rieke, E. L., Tracy, P. W., Morgan, C. L. S., & Honeycutt, C. W. (2020). Introducing the North American project to evaluate soil health measurements. *Agronomy Journal*, 112, 3195–3215. <https://doi.org/10.1002/agj2.20234>
- Nunes, M. R., Veum, K. S., Parker, P. A., Holan, S. H. J. P., Van Es, H. M., Wills, S. A., Seybold, C. A., & Karlen, D. L. (2024). SHAPEv1.0 scoring curves for dynamic soil properties. *Soil Science Society of America Journal*, 88(3), 858–875. <https://doi.org/10.1002/saj2.20668>
- Nunes, M. R., Veum, K. S., Parker, P. A., Holan, S. H., Karlen, D. L., Amsili, J. P., Van Es, H. M., Wills, S. A., Seybold, C. A., & Moorman, T. B. (2021). The soil health assessment protocol and evaluation applied to soil organic carbon. *Soil Science Society of America Journal*, 85, 1196–1213. <https://doi.org/10.1002/SAJ2.20244>
- Palusynska, A., Biecek, P., & Jiang, Y. (2022). Randomforestexplainer: Explaining and Visualizing Random Forests in Terms of Variable (R package version 0.10.1) [Computer software]. CRAN. <https://github.com/ModelOriented/randomForestExplainer/tree/v0.10.1>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Rieke, E. L., Bagnall, D. K., Morgan, C. L. S., Flynn, K. D., Howe, J. A., Greub, K. L. H., Mac Bean, G., Cappellazzi, S. B., Cope, M., Liptzin, D., Norris, C. E., Tracy, P. W., Aberle, E., Ashworth, A., Bañuelos Tavarez, O., Bary, A. I., Baumhardt, R. L., Borbón Gracia, A., Brainard, D. C., . . . Honeycutt, C. W. (2022). Evaluation of aggregate stability methods for soil health. *Geoderma*, 428, 116156. <https://doi.org/10.1016/j.geoderma.2022.116156>
- Six, J., & Paustian, K. (2014). Aggregate-associated soil organic matter as an ecosystem property and a measurement tool. *Soil Biology and Biochemistry*, 68, A4–A9. <https://doi.org/10.1016/j.soilbio.2013.06.014>
- Soil Survey Staff. (2022). *Kellogg soil survey laboratory methods manual* (Soil Survey Investigations Report No. 42). U.S. Department of Agriculture, Natural Resources Conservation Service.
- Topp, G. C., Reynolds, W. D., Cook, F. J., Kirby, J. M., & Carter, M. R. (1997). Physical attributes of soil quality. In E. G. Gregorich & M. R. Carter (Eds.), *Soil quality for crop production and ecosystem health* (Vol. 25, pp. 21–58). Elsevier.
- Van Eerd, L. L., Debruyne, A. H., Ouellette, L., Hooker, D. C., & Robinson, D. E. (2018). Quantitative and qualitative comparison of three wet aggregate stability methods using a long-term tillage system and crop rotation experiment. *Canadian Journal of Soil Science*, 98, 738–742. <https://doi.org/10.1139/cjss-2018-0101>
- Yoder, R. E. (1936). A direct method of aggregate analysis of soils and a study of the physical nature of erosion losses. *Agronomy Journal*, 28, 337–351. <https://doi.org/10.2136/sssaj1936.036159950B1720010046x>

How to cite this article: Aller, D., Amsili, J. P., & van Es, H. M. (2024). Cross-correlating soil aggregate stability methods to facilitate universal interpretation. *Agricultural & Environmental Letters*, 9, e20145. <https://doi.org/10.1002/ael2.20145>